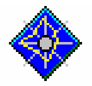
Voice Profile



The aim of this protocol is to objectify assessment of dysphonia and to help to the follow-up of vocal pathologies.

Principle

The subject pronounce a sustain « a » during a few seconds at usual intensity and pitch. The manipulator places an analysis zone on the most stable part of the signal. The program extracts from this observation zone indexes of voice quality like :

- mean, standard deviation, variation coefficient of F0, intensity and oral airflow.
- instability of vibration frequency : mean jitter , jitter factor, jitter ratio, Relative Average Perturbation (RAP)
- instability of vibration in amplitude : shimmer , Amplitude Perturbation Quotient, Shimmer factor
- signal/noise ratio
- glottic leakage index (explores presence or absence of « air leakage » through the glottis)

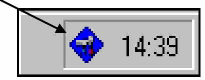
Preparation

Equipment

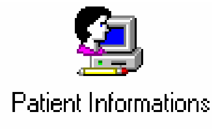
[EVA AERODYNAMIC]	[EVA ACOUSTIC]
Place a disinfected mask on the mouthpiece. Choose a mask that will fit well with the patient face.	Place the patient at 30 cm from the on stand microphone. Ensure yourself that he does not move. Plug the microphone jack in the INPUT 1-LEFT input.
Turn the selector INPUT 1-LEFT on MASK.	Turn the selector INPUT 1-LEFT on MICRO.

Software

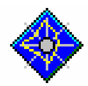
Launch the SESANE software by clicking this icon in Windows task bar.



In SESANE, enter the patient information :



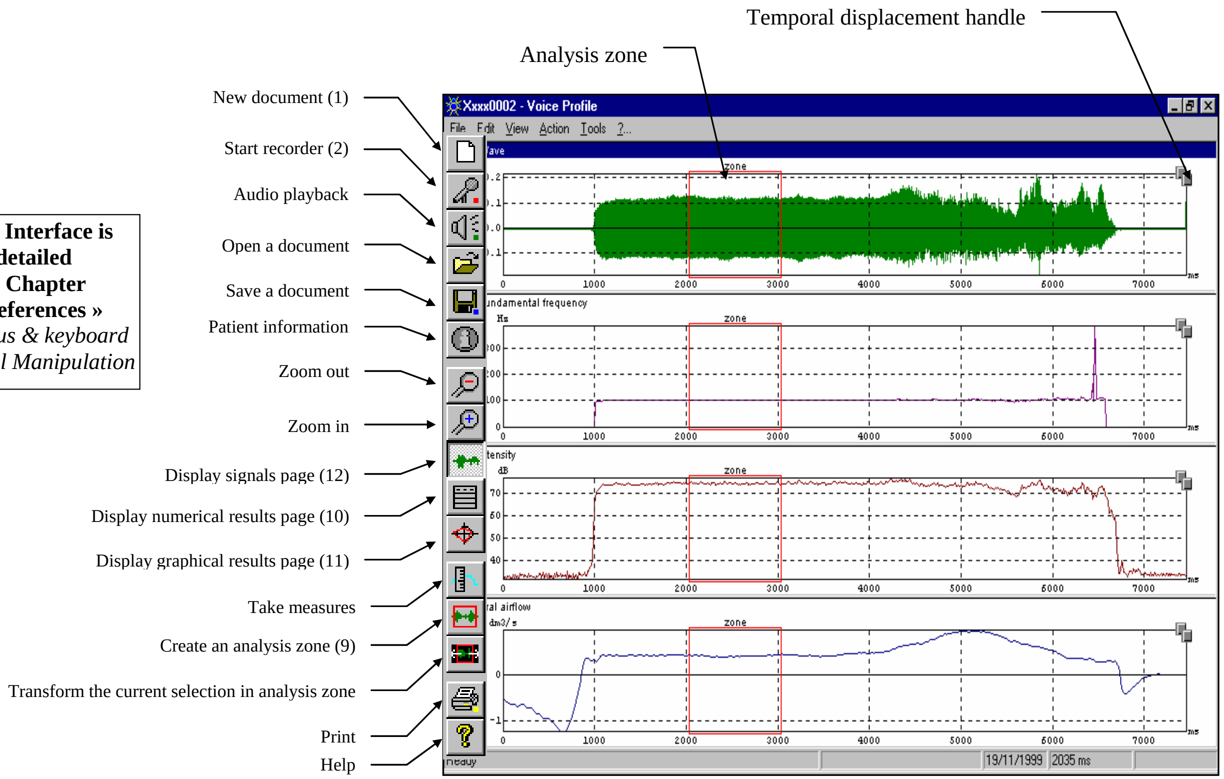
Then double click on this icon :



Using Voice Profile

Main window

User Interface is detailed in Chapter « References »
§ Menus & keyboard
§ Signal Manipulation



<u>Protocol</u>

(1) Create a new document if necessary
(2) Start real time display.

The recording control window appears.

[EVA AERODYNAMIC]
Move away the patient from the mouthpiece.
(3) Calibrate the sensors, wait three seconds. *The airflow level must be at zero.* Replace the patient in position.

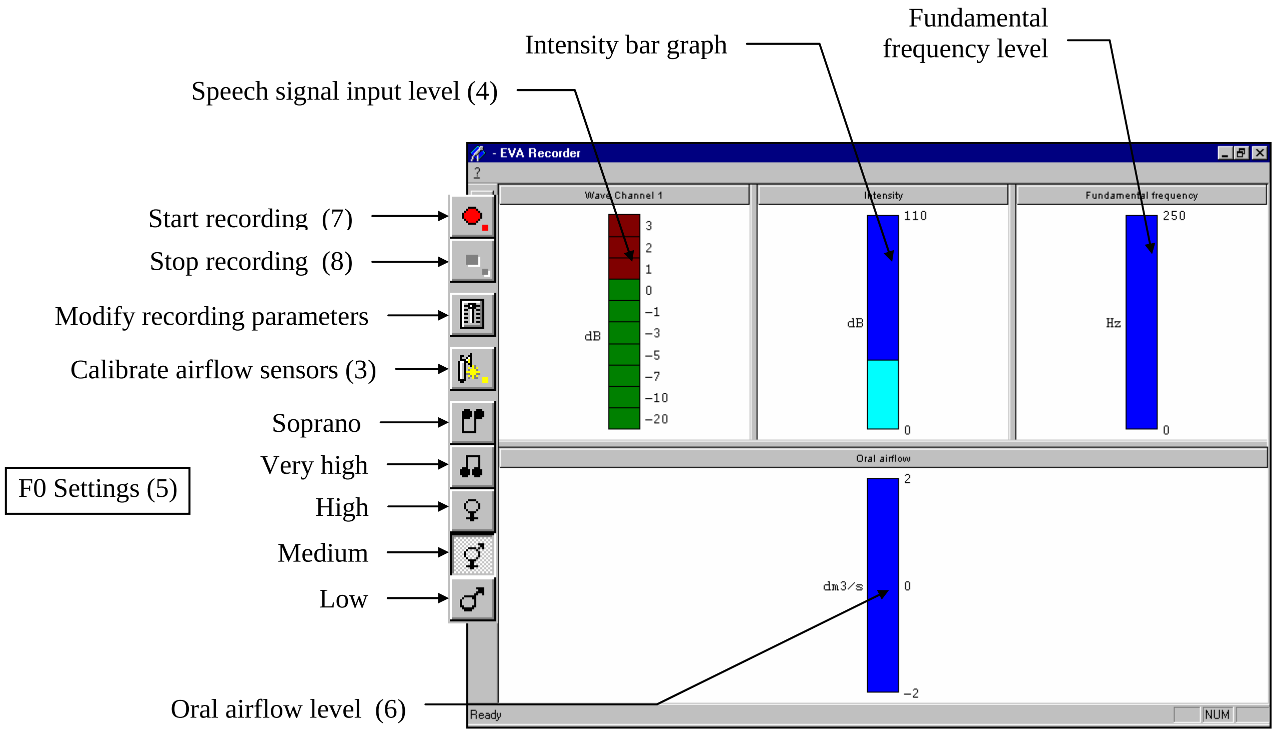
Make a try of a sustained « a » .
(4) Verify the recording level of the acoustic input vu-meter. Beware of not reaching +3 dB while recording. If necessary, adjust the volume button of INPUT 1-LEFT.
A low signal may indicate a bad position of the selector MASK - MICRO - LINE
(5) Set up the register of fundamental frequency depending on the patient tessitura.

[EVA AERODYNAMIC]
(6) Verify that the airflow level is correct. If necessary, verify that the patient is correctly pressed against the mask.

(7) Start recording.
The patient produces a sustained « a » at comfortable and usual pitch.
(8) Stop recording

The main window appears

Recording control window



(9) Create an analysis zone. *A zone cursor appears on the displayed signals.* Move this zone on the most stable part of the vocal emission. To do so, move the mouse pointer to the left edge of the zone cursor, click the left button and maintain it down, then move the analysis zone.

On this duration, computations are performed.

(10) Visualize the numerical results.
(11) Visualize the graphical results.

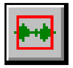
Explore if necessary other parts of the patient production.

Save the document.
Print document.

Measurement

Voice quality assessment is performed on an observation window chosen on the recorded signal.

Observation zone creation

To create this analysis zone click on the icon : 

A zone cursor appears on the displayed signals.

The zone duration is fixed. But it can be changed by using the Options (Menu Tools | Options) or by calling the object Menu | Properties (see after).

Observation zone placement

You have different ways to place this zone :

- on the most stable part of the vocal emission
- at a fixed time (ex : 200 ms) after the voiced attack.

To move it :

1. place the mouse pointer near the left edge of the zone cursor,
2. click the left mouse button and maintain it down,
3. move the analysis zone,
4. release

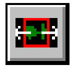
Editing observation zone

You can replace at any time this zone on another part of the recorded signal. All the analysis will be computed again.

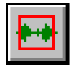
To define manually a zone :

1. select a part of signal (place the mouse pointer at the beginning, press and hold down the Shift key + left mouse button, move the mouse until the end and release).
   You can control the length of your selection by watching in the status bar.

Selection length

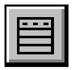
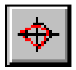
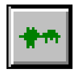
2. click on the icon of zone transformation : 

The analysis zone is enlarged to the size of the selected zone, and the statistics are immediately performed.

To get back to the standard zone, click on the icon : 

Results display

Results are displayed in separate pages : a statistics page which contains numerical and graphical results, and a radar page which summarise all theses results compared with the normality. To switch between these pages use these icons :

Statistics :  Radar :  Signals : 

You can also change the current page

- by selecting the menu « View | Switch to xxx page »
- OR by using the keyboard accelerators « 1 », « 2 », « 3 »

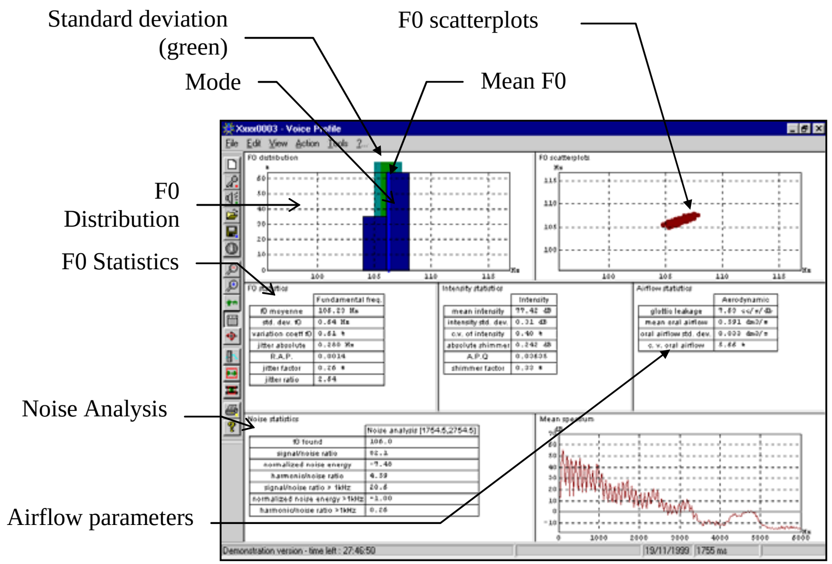
Statistics

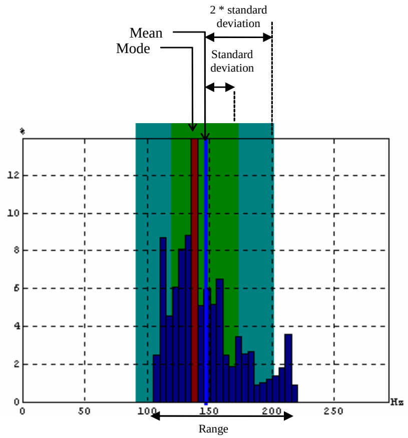
The program displays the mean, standard deviation, and coefficient of variation in percentage for each type of recorded signal.

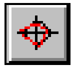
Several instability indexes are proposed :

- JITTER computed on F0 (measures frequency differences between vocal cycles)
- SHIMMER on amplitude (measures amplitude variations between vocal cycles)
- aerodynamic measures which explore the presence or absence of airflow leakage trough the glottis during the phonation
- signal to noise ratio
- spectral analysis

Details can be read in the Technical Note (see after).

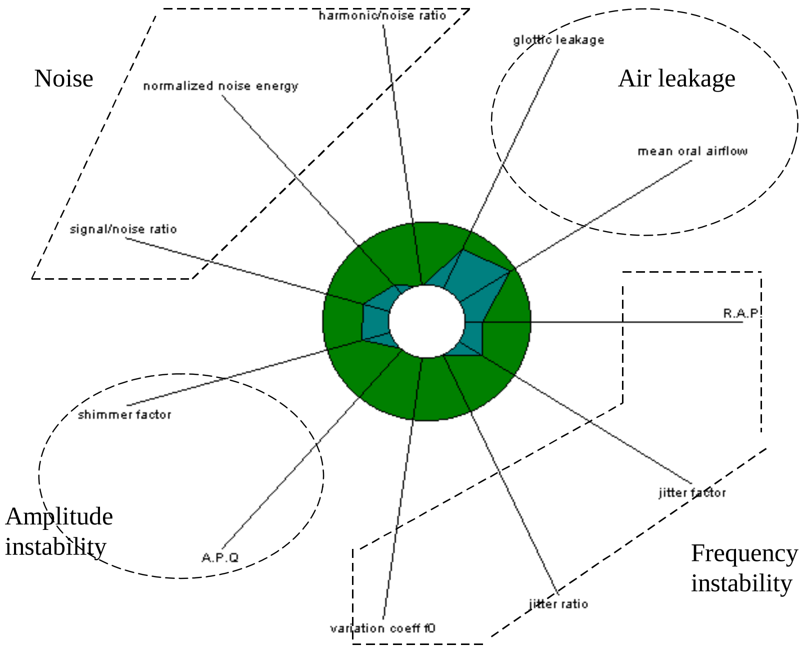




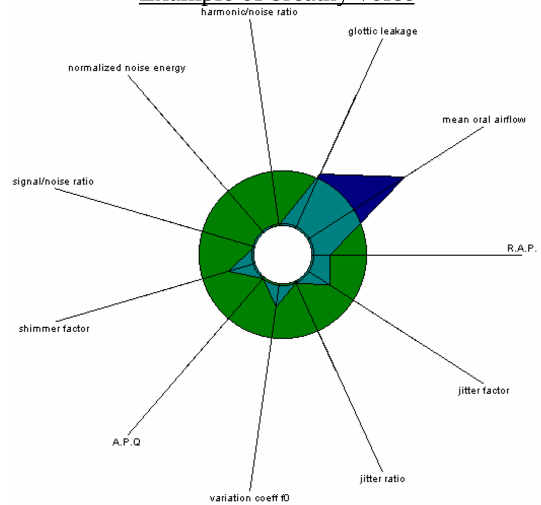
Radar display 

This representation allows to visualise in a very synthetic way a qualitative image of the voice parameters. It is a radar presentation where each axis explores a dimension of the voice. The circle shows the limits obtained with « normal » patients. If a measure exceeds this limit, it will be displayed outside the circle boundary. The values of these limits can be changed by using the Options (Menu Tools | Options) or by calling the object Menu | Properties (see after).

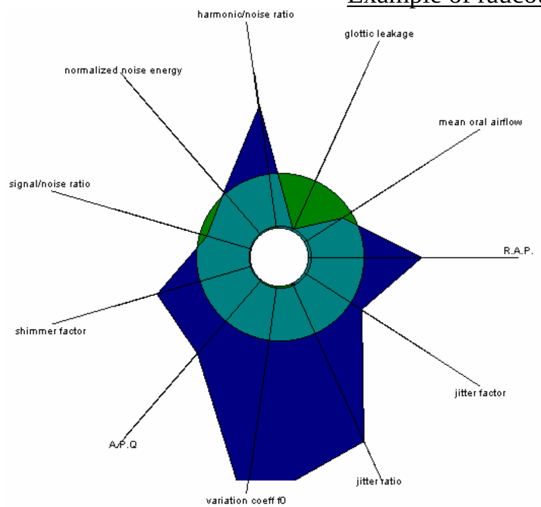
Example of normal voice



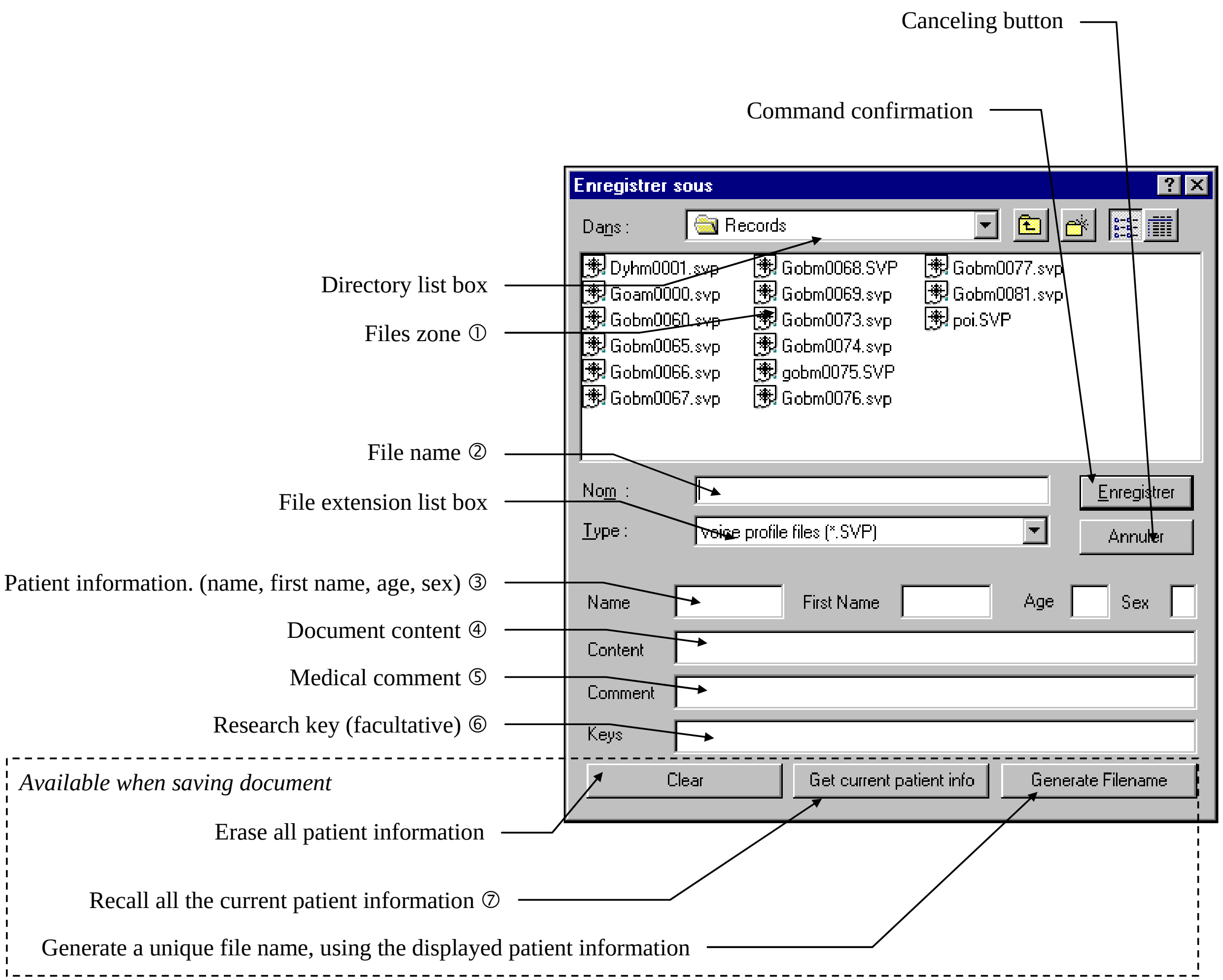
Example of breathy voice

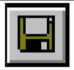


Example of raucous voice



Data Management





Save a document

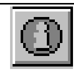
Method 1 : Click on ⑦. *The current patient information appear in the fields ③, ④, ⑤, ⑥. A unique filename is automatically generated in ②.* Confirm by clicking on ⑨.

Method 2 : Enter manually the patient information in the fields ③, ④, ⑤, ⑥. Click on ⑧. *A unique filename appears in ②.* Confirm the saving by clicking on ⑨.

Method 3 : Enter manually the patient information in the fields ③, ④, ⑤, ⑥. Enter a file name in ①. Confirm the saving by clicking on ⑨.

Open a document

Select a document in ① by a single click with the left mouse button.. *The file name appears in ② with its information as well in ③, ④, ⑤ ⑥.* Confirm your choice by clicking on ⑨.

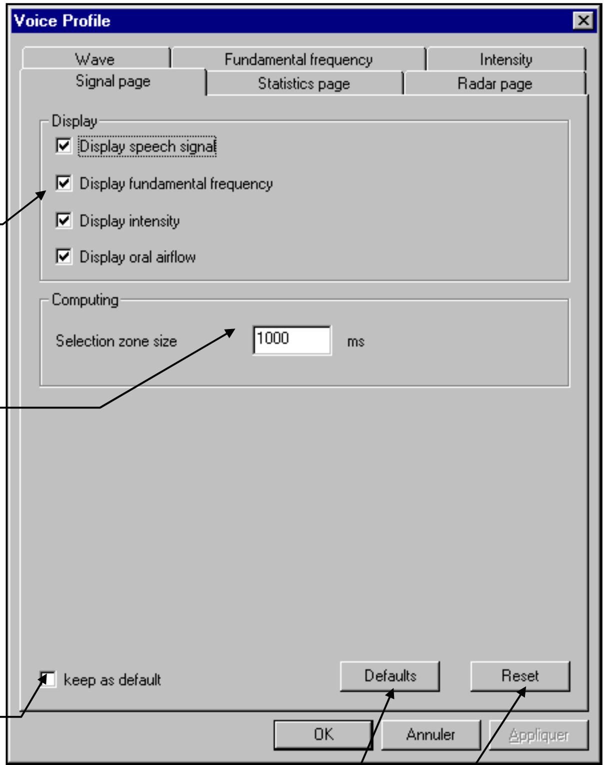


To obtain information about the current document

Options

You can modify the program parameters by selecting the menu « Tools | Options » or by typing the « O » key.

Signal Page Options



Use these checkbox if you want to show / hide some parameters

Change the selection zone size

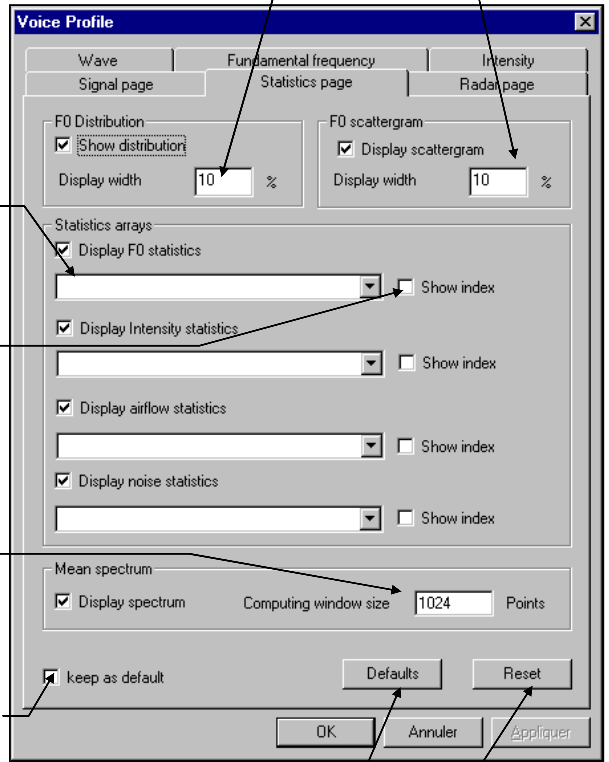
Set the changes as new program default

Retrieve the current program defaults

Retrieve the factory defaults

Statistics page options

Display width around the mean F0



<u>To modify the statistics arrays :</u>

① Select the index in the list box you want to show / hide

② Check / Uncheck the index

Size of the successively summed spectra

Set the changes as new program default

Retrieve the current program defaults

Retrieve the factory defaults

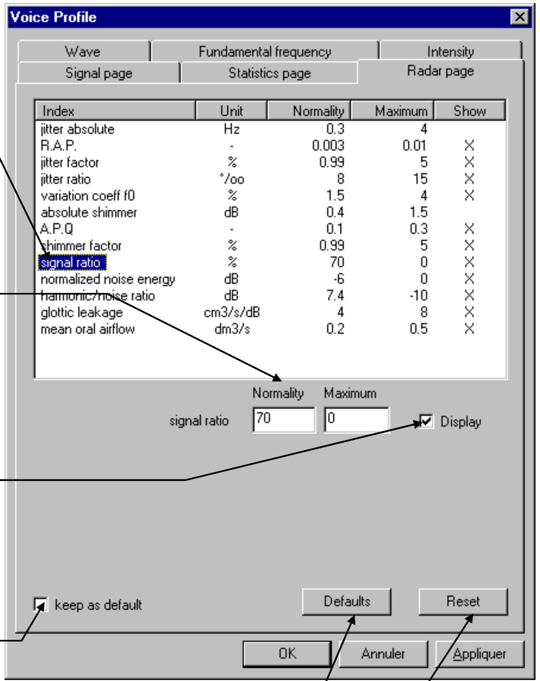
Radar Page Options

This option box allows to modify the normality values used in the radar representation. To do so, you must :

① Select the index you want to modify, by clicking on its name.

②The corresponding normality and its maximum value displayed is reported in theses text fields. Modify them, if you wish, the new values are immediately updated in the array.

③You may also choose to show or hide this index in the radar representation by selecting this check box.



Set the changes as new program default

Retrieve the current program defaults

Retrieve the factory defaults

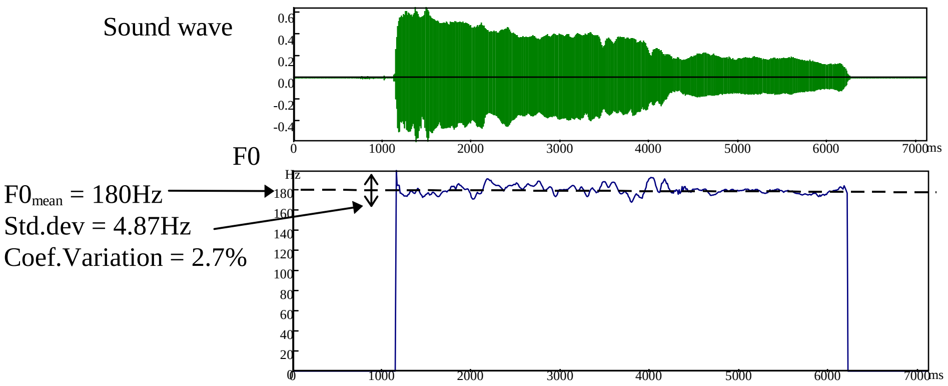
Technical Notes

Fundamental frequency statistics (F0)

These statistics explore the frequency of the vocal folds vibration.

(1) F0 Mean (Hz)
(2) F0 standard deviation
(3) Coefficient of variation of F0 (%)
(4) Mean Jitter (Hz)
(5) Relative Average Perturbation
(6) Jitter factor (%)
(7) Jitter ratio (‰)

	fundamental frequency
mean	179.97 Hz
std dev	4.87 Hz
c.v	2.71 %
CTC perturbation	0.677 Hz (jitter)
average perturbation	0.0015 (RAP)
percentage	0.38 % (jitter factor)
other	3.76 (jitter ratio)

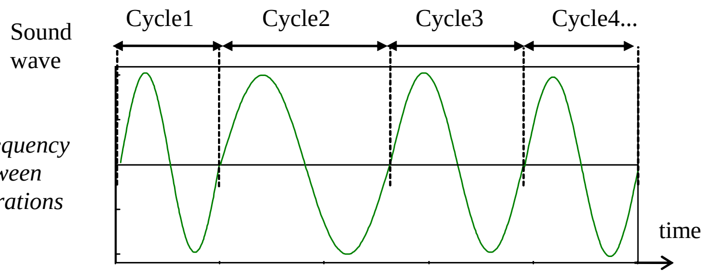


(1) **F0 mean** provides a global measure of pitch (high, medium, low...)

(2) Frequencies variations along the time can be seen as middle term instability of F0. It can be measured with the **F0 standard deviation,** which is relative to the scope in Hz of F0 variations around the average value.

(3) The **coefficient of variation** is interesting to relativize the standard deviation by comparing to the mean. It can be viewed as the scope in % of F0 variations around the average value. For example, with a 4.9 Hz standard deviation and a 180Hz average F0, the value of the coefficient of variation is 2.7%, which is important. With the same standard deviation but a 500 Hz mean F0, the value of the coefficient of variation is 0.98%, which is better. Finally, the coefficient of variation is the best index to explore the middle term instability of F0. It can be high with phenomena like pathological vibrato, quavering, neurological instability.

(4) The short term instability of F0 can be visible by comparing frequency variations between consecutive vibrations. **Mean JITTER** is the average value of the frequency differences between two consecutive vibrations. These variations are computed period per period, very precisely.



(5) **RAP** measures frequency variations but take into account three successive periods.

(6) **JITTER factor** is useful to relativize the mean jitter by dividing by the average value of F0. For example, with a 0.677 Hz mean jitter and a 180Hz average, we find a 0.38% jitter factor, low value. With the same mean jitter but a 80 Hz mean F0, we find a 0.84% jitter factor, more important value. Finally, the jitter factor is the best index to explore the short term instability of F0. It is well correlated with roughness.

(7) **JITTER ratio** measures the average difference of periods between two successive vibrations and relativize this value with the mean period.

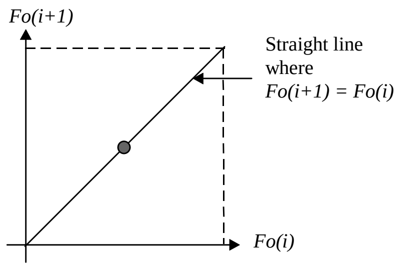
Remark : a high Jitter is always associated to a high coefficient of variation. The opposite is false. Some small successive F0 variations can furnish a low jitter. But if these variations are continuously positive (or negative), they increase the F0 scope and finally furnish a high coefficient of variation (typical example : vibrato).

About F0 phases diagram

F0 phases diagram is a scatterplot where the coordinates of a point are :
$x = Fo_i$ (Fo of the i^{th} cycle) , *$y = Fo_{i+1}$ (Fo of the i+1th cycle)*

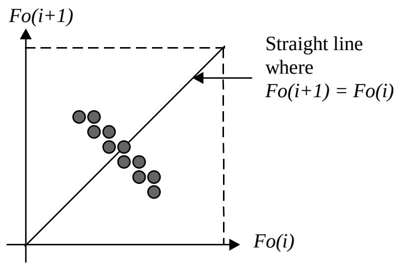
If the fundamental frequency is very stable (*Fo* constant), only a point appears on the scatterplot :

Scatterplot where Fo is very stable



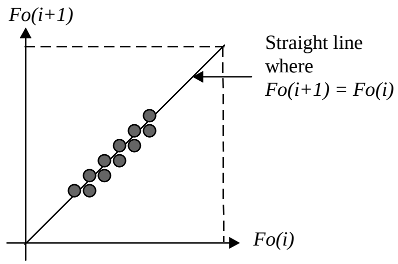
If the fundamental frequency is not stable in a short term (*Fo_i* very different from *Fo_{i+1}*), the points are distributed in a up-left to down-right line on the scatterplot :

Scatterplot with important jitter (Fo_i very different from Fo_{i+1})



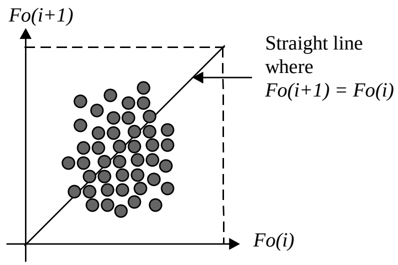
If the fundamental frequency is stable in a short term (low jitter) but not stable in a long term (ex : vibrato), the points are distributed in a down-left to up-right line on the scatterplot :

Scatterplot with important coefficient of variation (ex : vibrato)



If the fundamental frequency is never stable in a short and long term, the points are scattered :

Scatterplot with non stable Fo



Intensity/amplitude statistics

These statistics explore the sound amplitude and intensity.

(1) Mean Intensity (dB)

(2) Standard deviation of intensity (dB)

(3) Coefficient of variation of intensity (%)

(4) Mean Shimmer (dB)

(5) Average Perturbation Quotient (ratio)

(6) Shimmer factor (%)

	intensity
mean	86.67 dB
std dev	0.96 dB
c.v	1.11 %
CTC perturbation	0.163 dB (shimmer)
average perturbation	0.04739 (APQ)
percentage	0.19 % (Shimmer factor)
other	

(1) **The mean intensity** provides a global measure of loudness (poor, medium, strong...)

(2) The **standard deviation** gives information on the scope in dB about intensity variation around the average value.

(3) The **coefficient of variation** is necessary to relativize the standard deviation according to the mean intensity.

The previous parameters are based on the intensity which is measured with an 10 ms integration time, independently from the F0.

The following parameters are computed on the amplitude of each vibration.

(4) The short term instability of amplitude can be visible by comparing amplitude variations between consecutive vibrations. **Mean SHIMMER** is the average value of the amplitude ratios between two consecutive vibrations. This value is generally converted in dB. These variations are computed period per period, very precisely.

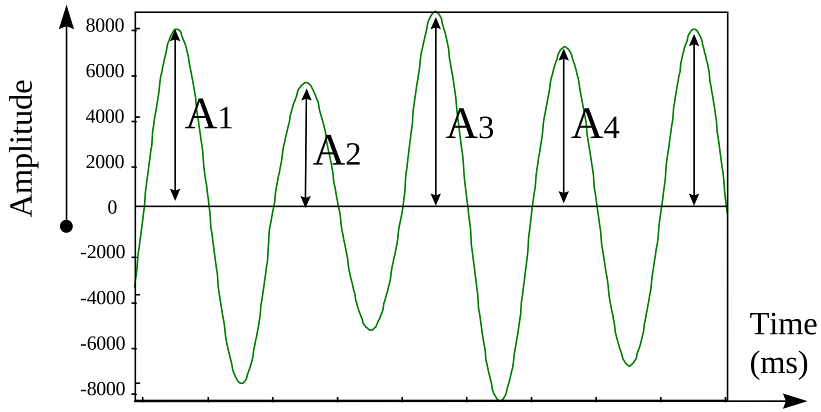
(5) **APQ** measures amplitude variations but take into account 11 successive periods.

(6) **SHIMMER factor** is useful to relativize the mean shimmer according to the mean amplitude.

Remark :

Mean, standard deviation and coefficient of variation in intensity are measured on the intensity curve which is independent from the recording level.

Instability parameters of amplitude (shimmer, APQ) are computed on the sound wave. Changing the input level during the recording time can spoil the results.



Noise statistics

Theses statistics explore the presence (or absence) of noise during phonation. This noise is due to turbulent flows caused by a bad closure of the vocal folds. We call it « additive noise ». A very unstable vibration can also be detected with this analysis : the difference of the f0 from its mean value is detected as a modulation noise.

This analysis is performed with temporal and spectral methods.

Harmonic/Noise Ratio is computed using a temporal method as described by Yumoto et al. (E.Yumoto, W.Gould, (1982) « Harmonics-to-noise ratio as an index of the degree of hoarseness », J.A.S.A, 71(6), 1544-1550) :
25 consecutive pitch periods of the sound wave are averaged ; H is the energy of the averaged wave form, while N is the mean energy of the differences between the individual periods and the averaged wave form. The ratio H/N, is computed along the time and the value is given in dB.

Signal Ratio is computed using a spectral method as described by Hiraoka et al. (N. Hiraoka, Y. Kitazoe, H.Ueta, (1984), « Harmonic-intensity analysis of normal and hoarse voices », J.A.S.A, 76(6), 1648-1651) :
we consider that the speech signal is the sum of two components :

- a periodic component (F0 & harmonics)
- a noise component (the rest)

The mathematical sum of theses two components is considered as the full signal.

SR Relative Signal Intensity = Energy (*F0 + Hi*) / Total energy (%)
signal ratio is computed on the whole spectrum.
signal ratio [f > 1kHz] is computed on the spectrum part above 1 kHz.

Normalized Noise Energy is computed using a spectral method as described by Kasuya et al. (H.Kasuya, S.Ogawa, K. Mashima, S.Ebihara, (1986), « Normalized noise energy as an acoustic measure to evaluate pathologic voice », J.A.S.A, 80(5), 1329-1334 :
the method is very similar to the previous one except in the final formula where

NNE Normalised Noise Energy = 10. Log [Energy(noise) / Total Energy] (dB)

Aerodynamic Measurements

These statistics explore the presence or absence of airflow leakage trough the glottis during the phonation.

$$\text{Glottic leakage} = \frac{OAF_{mean}}{Intensity_{mean}} \quad \text{in (cm3/s/)/dB}$$

This index evaluates the quantity of air necessary to produce a decibel, providing an estimate of glottic inefficiency.

Normality

Parameters displayed in radar representation :

Parameter	Unit	boundary « normality » / pathologic	Upper pathologic limit displayed	correlation with pathology*
Absolute Jitter	Hz	0.3	4	+
RAP	none	0.003	0.01	+
Jitter Factor	%	0.99	5	+
Jitter Ratio	‰	8	15	+
c.v. F0	%	1.5	4	+
Shimmer	dB	0.4	1.5	+
APQ	none	0.1	0.3	+
Shimmer factor	%	0.99	5	+
Signal Ratio	%	70	0	-
H/N ratio	dB	7.4	-10	-
NNE	dB	-6	0	+
Glottic leakage	cm3/s/dB	4	8	+
Oral Airflow	dm3/s	0.2	0.5	+

* The sign (+) means that higher is the parameter, higher the pathological aspect can be considered.
The sign (-) means that lower is the parameter, higher the pathological aspect can be considered.

Indexes Formulas

Fundamental Frequency indexes

Mean	mean (Hz) = $\frac{1}{N}\sum_{i=1}^{N} F0_i$
std dev	standard deviation (Hz)= $\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(F0_i - F0_{moy}\right)^2}$
c.v	coefficient of variation (%) = $100 \times \frac{\text{standard deviation}}{mean}$
Cycle to Cycle perturbation	Mean Jitter (Hz) = $\frac{1}{N-1}\sum_{i=1}^{N-1}\left\|F0_i - F0_{i+1}\right\|$
Average perturbation	Relative Average Perturbation RAP= $\frac{\frac{1}{N-2}\sum_{i=2}^{N-1}\left\|\frac{T0_{i-1}+T0_i+T0_{i+1}}{3} - T0_i\right\|}{T0_{moy}}$
percentage	Jitter factor (%) = $100 \times \frac{\text{mean jitter (Hz)}}{F0_{mean}}$
	Jitter ratio (‰) = $1000 \times \frac{\text{mean jitter (ms)}}{T0_{meany}}$

Intensity/amplitude indexes

Mean	mean (dB SPL) = $\frac{1}{N}\sum_{i=1}^{N} Intensity_i$
std dev	standard dev. (dB SPL) = $\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(Intensity_i - Intensity_{moy}\right)^2}$
c.v	coefficient of variation (%) = $100 \times \frac{\text{standard deviation}}{mean}$
Cycle to Cycle perturbation	shimmer$_{\text{dB}}$ = $\frac{1}{N-1}\sum_{i=1}^{N-1}\left\|20.\log\left(\frac{A_i}{A_{i+1}}\right)\right\|$
Average perturbation	Amplitude Perturbation Quotient APQ = $\frac{\frac{1}{N-10}\sum_{i=6}^{N-5}\left\|\frac{A_{i-5}+\ldots+A_i+\ldots+A_{i+5}}{11} - A_i\right\|}{A_{mean}}$
percentage	Shimmer factor = $100 \times \frac{\text{shimmer}}{20.\log\left(A_{mean}\right)}$

Aerodynamic indexes

Glottic leakage = $\frac{OAF_{mean}}{Intensity_{mean}}$ in (cm3/s/)/dB